ABSTRACT:

A computer program search engine is disclosed which identifies descriptions from a subject database that are conceptually similar to the target input string. The matching is based on a fuzzy logic correlation between the input terms, supplemented by semantically related terms, and each description in the subject database. The semantic relationships and contextual significance of the subject database's core vocabulary is initially generated off-line by manually applying a set of templates to a statistical analysis and sampling of the terminology usage in the descriptions. This data is stored in a look-up table that is used to expand an inputted target set of words and identify the relative importance of each term. Each description is then compared to the expanded target set with each word match extended by its significance factor. The conceptual similarity between a description from the database and an input string is expressed by the sum of all extended matches. Users are presented with the matches in descending order of similarity for entries with positive totals.